# From Frequent Hieroglyphic Signs to the Fringes of the Repertoire History and Prospects of the TSL <br> Philipp Seyr (ULiège) <br> Jorke Grotenhuis (UCLA Berkeley), Stéphane Polis (F.R.S.-FNRS; ULiège) 

LIĖGE

## Outline of the talk

- History of the project


## Outline of the talk

- History of the project
- Goal and data model


## Outline of the talk

- History of the project
- Goal and data model
- Recent developments
- Hieroglyphic repertoire expansion in Unicode
- Well known but not recorded hieroglyphs
- Local variations

History of the project

## History of the project

-2013


The Manuel de Codage Encoding of Hieroglyphs Impedes Development of Corpora*


Dictionnaire hiéroglyphique, inventaire des hiéroglyphes et Unicode

## History of the project

- 2013


## Réviser le codage de l'égyptien ancien

Vers un répertoire partagé des signes hiéroglyphiques
Stéphane Polis ${ }^{1}$, Serge Rosmorduc ${ }^{2}$


## History of the project

- 2013
- 2015


St. Polis \& S. Rosmorduc
A shared repository of hieroglyphic signs: The Thot Sign List (TSL)

## History of the project

- 2013
- 2015


28. January 2015, No. 03/2013

2015 Anneliese Maier Research Award winners announced

## Jean Winand

Mit Computerlinguistik zum ägyptischen Wörterbuch
Der belgische Ägyptologe Jean Winand zählt zu den führenden Experten für die ägyptische Sprache. Mithilfe der Computerlinguistik begann er 2006 mit dem Ramses-Projekt die Arbeit an einer Textdatenbank, welche die jüngeren Phasen der ägyptische Sprachgeschichte abdeckt, und führte damit moderne linguistische Methoden in die Ägyptologie ein. Hoch angesehen ist Winand auch als Co-Autor einer mittelägyptischen Grammatik und für seine Arbeiten zu regionalen sprachlichen Unterschieden und zur Kodierung von Hieroglyphen. In Deutschland will Winand gemeinsam mit seinem Heidelberger Partner Joachim Friedrich Quack die Erkenntnisse aus seinem Ramses-Projekt mit Arbeiten aus Deutschland, wie der der Berliner Akademie der Wissenschaften an einem Thesaurus Linguae Aegyptiae, zusammenführen und ein umfassendes modernes Wörterbuch des Ägyptischen schaffen

Gastinstitut: Universität Heidelberg, Ägyptologisches Institut
Gastgeber: Prof. Dr. Joachim-Friedrich Quack
Prof. Dr. Jean Winand
wurde 1962 in Belgien geboren. Er ist Professor an der Université de Liège, sowie an der l'Université libre de Bruxelles, beides Belgien. Zudem ist er Mitglied in mehreren nationalen und internationalen Gremien der Ägyptologie. In seinen Arbeiten konzentriert er sich auf die Textanalyse und Lexikographie der ägyptischen Sprache und Schrift.


Foto: Université de Liègel Michel Houet

## History of the project

- 2013
- 2015



## History of the project

- 2013
- 2015-2016


HOME

Q Filter concepts in curre 0


* Thot Collections
> * Dates and dating systems
> * Language
> * Material
> * Museums and private collection
> * Scripts
* Technique of inscription
> * Text content


## Thesauri

Using the definition set by the International Standard Organisation, thesauri are controlled and structured vocabulary in which concepts are represented by terms, organised so that relationships between concepts are made explicit, and preferred terms are accompanied by lead-in entries for synonyms or quasi-synonyms.' In Thot, each concept is represented by one or several terms expressed in Arabic, English, French and German, and identified by a unique identifier (the 'thot-number'). It should also be noted that a supplementary 'language" has been included, the so-called 'Thot-xml', in order to provide terms to be used as values of XML attributes in digital documents, such as TEI files.

Thesauri covered by the project
As a first approach, the Thot project aims at compiling a wide range of thesauri relating to artifactual and textual metadata. This will consist in quite small thesauri made of a couple of terms, as well as rather big ones populated by several hundreds of concepts. Subject headings pertaining to general and abstract concepts will be included in a second stage.

## History of the project

- 2013
- 2015-2016


Thot Sign List
Documents Sources Signs Tokens

## History of the project

- 2013
- 2015-2016
- 2016-2017



## History of the project

## - 2013

- 2015-2016
- 2016-2017



## History of the project

## - 2013

- 2015-2016
- 2016-2017



## History of the project

## - 2013

- 2015-2016
- 2016-2017



## History of the project

- 2013
- 2015-2016
- 2016-2017
- 2016-2019


History of the project

- 2013
- 2015-2016
- 2016-2017
- 2016-2019



## History of the project

- 2013
- 2015-2016
- 2016-2017
- 2016-2019
- 2019-today



## The Thot Sign List (TSL)

## An open digital repertoire of hieroglyphic signs*

# Stéphane Polis (F.R.S.-FNRS / ULiège), Luc Desert (ULiège), Peter Dils (SAWL), Jorke 

 Grotenhuis (ULiège), Vincent Razanajao (UBordeaux-Montaigne), Tonio Sebastian Richter (BBAW), Serge Rosmorduc (CNAM), Simon D. Schweitzer (BBAW), Daniel A. Werning (BBAW, HUBerlin), Jean Winand (ULiège)In THIS PAPER, we introduce the Thot Sign List (TSL), an online digital repertoire of hieroglyphic signs (http://thotsignlist.org) that records the graphemes attested in the Ancient Egyptian hieroglyphic texts, with special attention to their contextual functions and to their iconic variations across media and time. The paper is structured as follows. In Section 1, we present a brief history of the TSL and we specify its goals and targeted audience. In Section 2, we introduce the data model, which meets the requirements formulated in Meeks (2013) and Polis \& Rosmorduc (2013), and discuss the use of shared thesauri for the metadata (Thesauri and Ontologies for documenting Ancient Egyptian resources; https://thot.philo.ulg.ac.be). Section 3 describes the user interface and discusses how data can be browsed, searched, and visualized by users depending on their level of access. We describe the way credits and citation work for this database in Section 4, and we give a list of the collaborators. In the conclusions, we argue that such an online resource will immensely benefit from monitored crowdsourcing: Egyptologists all around the world can enrich this digital repertoire with new sources and examples of hieroglyphic signs.

## TSL - encoded data

TSL data (end 2021)


## TSL - encoded data

TSL data (end 2021)


## TSL - encoded data

TSL data (end 2021)


## Hieroglyphic signs

Filter: Parts of the human body [Gardiner D]
$\checkmark$ full quadrat $\checkmark$



| Functions Classes | Tokens | Cite as |  |
| :--- | :--- | :--- | :--- |
| Classifier (1 function(s)) | Semantic value |  |  |
| Phonetic value | movement by boat | Use |  |
| - | Semalar |  |  |
|  | to row, to convey by water |  |  |
| Logogram (1 function(s)) |  | Use |  |
| Phonetic value | regular |  |  |
| hnj |  |  |  |
| Phonemogram (1 function(s)) | Semantic value |  |  |
| Phonetic value | - | Use |  |
| hn |  | regular |  |
|  |  |  |  |

hnj

| Token | Source | Date |
| :---: | :---: | :---: |
| 5 | hnj jw(=j) hn.n(=j) wr.thr zpf <br> Row! And I rowed powerfully, said this man. | Kaiuhor Menkauhor |
| $\pi$ | SRAF <br> hn sw jm. yw 3h.t sqd sw jm.yw qbh.w <br> the ones in the horizon will row him, the ones in the cool waters will sail him | Unas |
| 5 |  <br> hn sw mw.t=f <br> Row him, his mother! | Unas |
| $9$ | nm p.t hnn ${ }^{c}$ s 3 h h hn dw3.t hnn ${ }^{c}$ wsjr <br> [...] who travels the sky with Orion and rows the Duat with Osiris. | Nemtiemsaf I Merenre I |
|  | $\begin{aligned} & \text { 周 } \\ & \text { hsfwjz.t=kn.tj.hm.w-sk.wjrhn.t=k} \\ & \text { As your crew of Imperishable Stars should not be barred from rowing you, [...] } \end{aligned}$ | Nemtiemsaf I Merenre I |
| $31$ | hft d (3)w hm n ntr pn šps rjr.t hnn.t=fm jp.t=frs.(y)t when the majesty of this noble god proceeds to do his water procession in his southern inner room, | Tuthmosis III Menkheperre (complete reign) |

Description: Two arms, holding an oar, in a rowing motion, below a cover.
Codes
$>$
Gardiner D33
Hieroglyphica D33
Jsesh D33
Unicode U+13099 凡

Bibliography
Borghouts, J. F. (2010), Egyptian: an introduction to the writing and language of the Middle Kingdom, 2 vols. Egyptologische Uitgaven 24, page 32, note 1
Gardiner, Alan Henderson (1957), Egyptian grammar being an introduction to the study of hieroglyphs, 3rd, revised edition, page 453
$>$ Credits $>$
Creator: L. Seelau
Editor(s): J. Grotenhuis, L
Seelau, S.D. Schweitzer

| TSL_1_2177_00 |  | TSL_1_2177_02 |
| :---: | :---: | :---: |
| Description > | Description > | Description > |
| Two arms, holding an oar, in a rowing motion, below a cover. | Two arms, holding an oar, in a rowing motion. | Two arms, holding an oar, in a rowing motion, under a cover resembling a mouth. |
|  | Basic Form: full quadrat |  |
| Basic Form: full quadrat |  | Basic Form: full quadrat |
| Tag: arms, oar | Type: composite | Tag: arms, mouth, oar |
| Type: composite | Components TSL_1_5126_00 d | Type: composite |
| Components TSL_1_5126_00 d |  | Components TSL_1_2020_00 o |
| Codes > | Hieroglyphica D33B | TSL_1_5126_00 |
| Gardiner D33 | Jsesh D33B |  |
| Hieroglyphica D33 <br> Jsesh D33 | Bibliography | Hieroglyphica D33C |
| Unicode U+13099 | Credits v | Jsesh D33C |
|  |  | Bibliography |
|  |  | Credits v |



Description: Two arms, holding an oar, in a rowing motion, below a cover.
Description:

- Codes
$\vee$

Functions Classes Tokens Cite as


## Search $>$

## OCriteria

|  | Function Type $\vee$ | $=$ | $\vee$ |  | $\vee$ |
| :--- | :--- | :--- | :--- | :--- | :--- |
|  | Function Type | $=$ |  | logogram | - |
| AND | Semantic Value | contains | to see | - |  |
|  |  |  |  |  |  |

## OTSL code (TSL1_)

## Reset Search

0
D4
An eye.
Parts of the human body [Gardiner D]

TSL1_2366

## $\infty$

D7A
An eye with an eyebrow.
Parts of the human body [Gardiner D]


Goal and data model

## Goal and data model

TSL is born out of the practical necessity of creating electronic corpora of hieroglyphic texts in a principled way and its first aim is

## to document the functions

attested for individual signs

## Goal and data model

TSL is born out of the practical necessity of creating electronic corpora of hieroglyphic texts in a principled way and its first aim is

## to document the functions

attested for individual signs

- References to ancient sources
- Description of the iconic features
- Structured repertoire


KIU 1730










# Recent developments 

Hieroglyphic repertoire expansion in Unicode
Well known but not recorded hieroglyphs
Local variations

## Hieroglyphic repertoire expansion in Unicode

- Currently 1072 Hieroglyphic signs in Unicode
- Primarily an expansion based on Graeco-Roman signs.
- Introduction of Control Characters (Accepted, and likely available in 2023).
- Insertion - Overlay - Mirroring - Rotation - Enclosure controls - Shading

- For more detail see Nederhof et al., Additional control characters for Ancient Egyptian hieroglyphic texts (https://www.unicode.org/L2/L2021/21208-egyptian-ctrl.pdf)


## Hieroglyphic repertoire expansion in Unicode

- Currently 1072 Hieroglyphic signs in Unicode
- Primarily an expansion based on Graeco-Roman signs
- Introduction of Control Characters (Accepted, and likely available in 2023)
- Expansion of Hieroglyphic repertoire
- Unicode project funded through a Google Research grant awarded to Dr. Deborah Anderson at UC Berkeley (Performed by Jorke Grotenhuis)


## Hieroglyphic repertoire expansion in Unicode

- Currently 1072 Hieroglyphic signs in Unicode
- Primarily an expansion based on Graeco-Roman signs
- Introduction of Control Characters (Accepted, and likely available in 2023)
- Expansion of Hieroglyphic repertoire
- Creation of a candidate list
- Two candidate lists
- Unicode candidate list: Inclusion based on data from modern and ancient sources: TLA, Ramses, TSL, Karnak database, Cauville, Kurth, Valeurs phonétiques (Daumas), Hannig, Dendera (Hallof), Athribis, Kom Ombo.


## Hieroglyphic repertoire expansion in Unicode

- Currently 1072 Hieroglyphic signs in Unicode
- Primarily an expansion based on Graeco-Roman signs
- Introduction of Control Characters (Accepted, and likely available in 2023)
- Expansion of Hieroglyphic repertoire
- Creation of a candidate list
- Two candidate lists
- Unicode candidate list
- Unicode core list: A curated list, where every sign is verified with an image (photo or facsimile of an Ancient source) and accompanied by a description and function.


## Hieroglyphic repertoire expansion in Unicode

- Currently 1072 Hieroglyphic signs in Unicode
- Primarily an expansion based on Graeco-Roman signs
- Introduction of Control Characters (Accepted, and likely available in 2023)
- Expansion of Hieroglyphic repertoire
- Creation of a candidate list
- Two candidate lists
- Unicode candidate list: +/- 4000 signs (including the 1072 existing signs)
- Unicode core list: +/- 3000 signs (including most of the 1072 existing signs).


## Hieroglyphic repertoire expansion in Unicode

- Currently 1072 Hieroglyphic signs in Unicode
- Primarily an expansion based on Graeco-Roman signs
- Introduction of Control Characters (Accepted, and likely available in 2023)
- Expansion of Hieroglyphic repertoire
- Creation of a candidate list
- Two candidate lists
- Unicode candidate list
- Unicode core list
- A windows capable font (Andrew Glass)


## Well known but not recorded hieroglyphs



## Well known but not recorded hieroglyphs




## Well known but not recorded hieroglyphs



KV 9 (Ramses V/VI)

## $\Omega \Omega \Omega 1 s \Omega$



## Well known but not recorded hieroglyphs



KV 9 (Ramses V/VI)

## $\sim \Omega \Omega \Omega$

## Well known but not recorded hieroglyphs



## Well known but not recorded signs



## Local variants



## Local variants





Description: A saharan helmeted guinea fowl (Numida m. meleagris), with a lappet at the front of its

## Functions Classes Tokens Cite as



| Description |
| :--- |
| Codes |
| Bibliography |
| Credits |

Description
Codes
Bibliography
Credits

## 术


Description: A saharan helmeted guinea fowl (Numida m. meleagris), with a lappet at the front of its $v$ Codes
breast.
Bibliography

Functions Classes Tokens Cite as


## www.thotsignlist.org

We are looking forward to future collaborations!

